DOCUMENT RESUME

ED 164 557 TM 007 402

AUTHOR	Rudman, Herbert C.
TITLE	The Superintendent and Testing: Implications for the Curriculum. NCME Measurement in Education. Volume 8, No. 4, Fall, 1977.
INSTITUTION	National Council on Measurement in Education, East Lansing, Mich.
PUB DATE	Oct 77
NOTE	9p.
AVAILABLE FROM	Dr. Irvin J. Lehmann, 202 S. Kedzie Hall, Michigan State University, East Lansing, Michigan 48824 ($1.50 ea., 25 or more, $0.75 ea.)
JOURNAL CIT	NCME Measurement in Education; v8 n4 F 1977
EDRS PRICE	MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS	Data Analysis; Declining Enrollment; *Educational Accountability; *Educational Legislation; Educational Problems; *Educational Trends; Elementary Secondary Education; Expenditures; State Standards; *Student Testing; *Testing Problems; Testing Programs; *Test Interpretation
IDENTIFIERS	Test Score Decline

ABSTRACT

This discussion of educational trends of interest to school superintendents focuses on three issues: (1) the general decline of birth rates, school enrollments, staff, facilities, and expenditures for instruction; (2) the increase in mandated assessments and programs brought on by declining academic performance and political influences; (3) the care that must be taken in using and interpreting test data. The author suggests three variables that are important in assessing the effectiveness of teachers or school programs: the students' characteristics; the characteristics of their homes and communities; and those school-related factors which may affect test scores. School staff must recognize the fact, he adds, that the school has little or no control over some factors which contribute to student achievement. (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Volume 8, No. 4 NCME Fall, 1977

measurement in education

A SERIES OF SPECIAL REPORTS OF THE NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

I. J. Lehmann
NCME

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

ED164557

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

IM002 402

The Superintendent and Testing: Implications for the Curriculum

Herbert C. Rudman
Michigan State University

ABOUT THIS ISSUE

Dr. Herbert C. Rudman

As mentioned in earlier issues, ME is renewing its efforts to provide meaningful, practical information for the field-based professional. This audience includes, of course, those who function as the administrators of public school programs. They are responsible for the decision-making regarding curricular development, testing programs, accountability systems, etc., all of which ultimately affects how the schools perform their instructional functions.

In this provocative and enlightening article, Dr. Herbert C. Rudman examines some of the factors which impact upon school administrators, particularly superintendents, as they select and implement large-scale testing programs.

Dr. Rudman is widely published both in the professional literature and in trade books. He is particularly known as co-author of the 1964 and 1973 editions of the Stanford Achievement Test. He is currently Professor of Administration and Higher Education, Michigan State University, East Lansing, MI 48824.

PSR

The superintendent's world of today is a complex one that is in many ways much more difficult to confront and to reflect upon than the world we faced just two decades ago. Then we faced an abundance of children, a shortage of buildings and supplies, too many inefficiently sized and organized districts, a generally submissive citizenry, and a court system that did not concern itself with the day-to-day operations of the nation's schools. As many of us look back on those days they become the "golden years" of our professional lives. But you wouldn't have convinced me of it twenty and thirty years ago.

But today it's a new ball game. Now we face declining birth rates, declining enrollments (north of the Sun Belt, certainly), declining test scores, declining need for staff and facilities, declining proportion of expenditures and revenues available for instructional materials, and declining public support and confidence for public education. On the other hand we are confronted with increasing legislative demands for educational assessment and accountability, increasing bureaucracies and policies at state and national levels, increasing salary and benefit costs which are coupled with increasing utility costs and a general inflation, and increasing politicizing of testing coupled with an increasing intervention of court decisions into the administration of the schools.

The implications for curriculm? I believe that is easier to discuss the implications of the superintendent's contemporary world as it relates to his or her sanity!!

In short, were we to address even these ten factors — which certainly do not reflect the total concerns of the contemporary superintendent — the length of this paper would need to be extended for longer than space would permit. I have chosen, instead to address three major points that concern the superintendent and his/her professional staff:

1. The general decline of birth rates, enrollments, staff needs, proportion of expenditures for curriculum materials and the like.

2. The general increase in mandated assessments and programs brought on by a general decline in academic performance as evidenced by test scores and out-of-school behaviors.
3. The care we must take in using and interpreting the data we collect.

The Declines

Stanley E. Hecker and Fred Ignatovich of Michigan State University have developed an approach to analyzing live birth rates and in-and-out migration of states. They have developed a series of projections which fortell the future of schooling in Michigan. While I confess to ignorance of the specifics of their forecasting techniques, I am acquainted with the data.[1]

Figure 1 is a computer-drawn graph of the actual live births recorded in Michigan for the years, 1961 to 1976. While this chart shows a steady decline, it does show one particular spurt in birth rate for the years 1969 and 1970. While I have no explanation for this particular anomoly, it would make an interesting topic for a doctoral dissertation. I can just see the correlational studies sprouting between the quality of television programs in those two years and the birth rate.

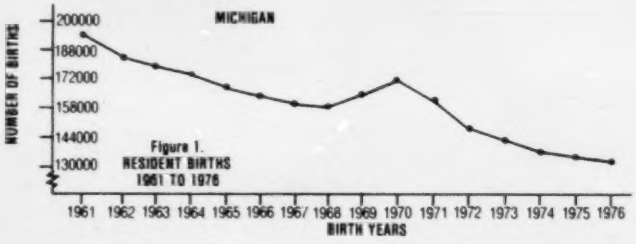

Figure 1.
RESIDENT BIRTHS
1961 TO 1976

The subsequent charts trace that bulge in births through the grades; but what we will see still is a steady decline in school enrollments that will reflect not only the birth rate, but movement from our state by the young child-bearing adults who for various reasons are following the sun and warmer climates a phenomenon being experienced in the Northeast.

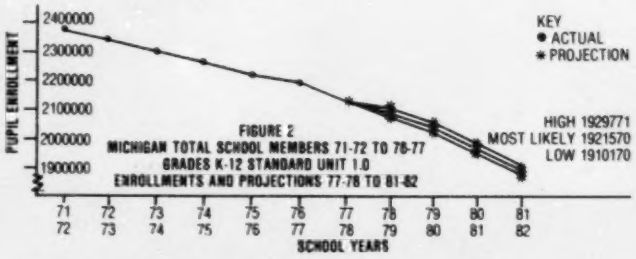

FIGURE 2
MICHIGAN TOTAL SCHOOL MEMBERS 71-72 TO 76-77
GRADES K-12 STANDARD UNIT 1.0
ENROLLMENTS AND PROJECTIONS 77-78 TO 81-82

Figure 2 illustrates the fact that Michigan, it is estimated, will lose over 220,574 children in her private and public elementray shcools, and will in all likelihood lose up to 9,174 teaching positions during the period 1976-1981.

Predicting anything is, of course, a risky business. The first quarter of this year, for example, indicates a projected 8% annual growth in Michigan's live births, but based on last year's total number of recorded live births in Michigan, this projected increase will still be very small. Last year, for the first time in some years, Michigan experienced an increase in its anticipated kindergarten enrollment which indicated that the excess expected enrollment was likely the result of an in-migration of young child-bearing adults. And this was probably due in part to an upswing in sales of automobiles and increased jobs in that industry. But because of a smaller base of child-bearing age adults, only the most optimistic among us dare talk of a baby boom in any of the non-sunbelt states. The conclusion I draw from all of this? Most of us can look forward to coping with the problems of fewer students, facilities, and staff for the rest of our professional lives.

Turning now to the curriculum and our ability to devote resources to it, the following data demonstrates a linkage between a decline in pupil enrollment and its impact upon our ability to focus our resources upon the curriculum. For all of our careers we have been faced with the problems that economic inflation brings to the administration of schools. But the period of the late '60's has introduced an added effect. Increased collective bargaining has brought about a dramatic upsurge of expenditures for salaries and other staff benefits. In many states, 80% or more of the

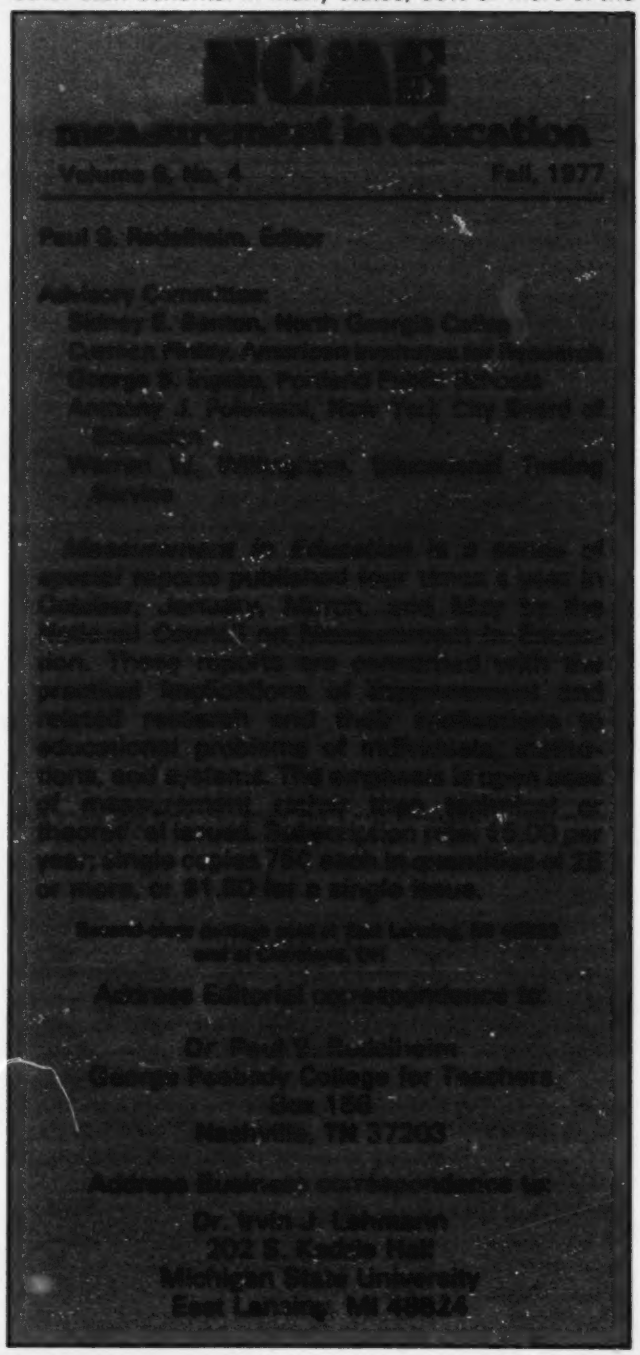
NCME
measurement in education
Volume 6, No. 4 Fall, 1977
Paul S. Nachtman, Editor
Advisory Committee
Sidney E. Barton, North Georgia College
Carmen Fusky, American Institutes for Research
George S. Ingle, Portland Public Schools
Anthony J. Pusateri, New York City Board of Education
William W. Turnbull, Educational Testing Service
Measurement in Education is a service of the National Council on Measurement in Education, Inc. and is published four times a year in October, January, March, and May by the National Council on Measurement in Education, Inc. Membership in the Council includes a subscription to Measurement in Education. Membership is open to persons professionally engaged in educational measurement or who are concerned with the problems of measurement in education. Membership dues are $20.00 per year; single copies $5.00 each; reprints of 25 or more, $1.00 for a single issue.
Address Editorial correspondence to:
Dr. Paul S. Nachtman
George Peabody College for Teachers
Box 158
Nashville, TN 37203
Address Business correspondence to:
Dr. Wren J. Lehmann
202 E. Kedzie Hall
Michigan State University
East Lansing, MI 48824

TABLE 1
A FIVE-YEAR ANALYSIS OF EXPENDITURES
FOR TEST BOOKS AND TEACHING SUPPLIES EXPRESSED
IN PROPORTION OF TOTAL OPERATING EXPENSE

Budget Items	Years 1971-72 $	%	1972-73 $	%	1973-74 $	%	1974-75 $	%	1975-76 $	%
TEXTBOOKS										
K - 6	6,581,027	.3	6,771,928	.3	7,253,601	.3	7,611,630	.2	7,140,966	.2
7 - 12	7,023,334	.4	7,526,115	.4	7,609,747	.3	8,217,534	.3	8,448,429	.3
Total	13,604,361	.7	14,298,043	.7	14,863,348	.6	15,829,164	.5	15,589,395	.5
TEACHING SUPPLIES										
K - 6	16,465,192	.8	17,680,579	.8	20,061,330	.8	21,978,769	.8	21,507,782	.7
7 - 12	17,794,598	.9	20,705,305	1.0	24,617,405	1.0	28,710,720	1.1	28,866,132	1.0
Total	34,259,790	1.7	38,385,884	1.8	44,678,735	1.8	50,689,489	1.9	50,373,914	1.7
TOTAL OPERATING EXPENSES	1,947,259,794		2,099,577,184		2,286,588,623		2,564,467,258		2,793,105,142	

total operating budget is allocated to instructional and non-instructional salaries, and more than two-thirds of that total operating budget is allocated for instructional salaries. Now when we couple this with a dramatic increase in utility costs and an overall inflation in the costs of all types of materials and services, there is little room left for the purchase of instructional materials; a not insignificant concern we must all share. Table 1 illustrates just how small a proportion of school operating budgets is spent on textbooks and other teaching supplies.[2] If we note the expenditures for the elementary schools we can see increasing funds spent through 1974 with a corresponding drop noted in 1975-76. Part of this increase can be explained by increased costs of these supplies, but the proportion of the budget allocated remains stable for the most of that five-year period. Secondary school expenditures show an increase in actual dollar amounts, but a relatively stable proportion allocated to the purchase of teaching materials.

Undoubtedly the decrease in elementary school enrollment is playing an increasing role in the funds spent on instructional materials at that level and a similar phenomenon will soon catch up with the secondary schools. But there is considerable reason to believe that continuing pressures for staff salaries and benefits will erode even further the proportion of total operating expenses available for the purchase and use of instructional materials.

One last observation: Michigan auditing procedures indicate the category under which local school districts must report their expenditures for standardized tests, but when the state summarizes total state expenditures, the actual monies spent for standardized testing is subsumed under the line item, "Teaching Supplies." I have drawn a sample of nine school districts that range in size from large city to rural. From this sample we can ascertain the actual percent of the "Teaching Supplies" line item spent for standardized tests and can thus estimate what proportion of the state summary of expenditures is used for testing.[3] Table 2 indicates the results of that sampling. As we can see, only a fraction of one percent is spent by local districts in Michigan for standardized tests. (This figure represents only hard-money expenditures; no compensatory aid — state or federal — is included).

Someone not truly familiar with the headaches of the superintendency might think that with fewer children in the schools the job of school administration is less of a problem than it was twenty or thirty years ago when we faced the prospect of supplying an adequate education for a burdgeoning school population. But as we know, that just isn't the case. In many ways it was much easier to mount a building program than it is to convince a community that they no longer can efficently keep a school building open. It was easier to justify a millage election for updating and modernizing a curriculum than it is to justify a millage increase when we know that only one half of a penny for every dollar we collect will be spent on the purchase of new instructional textbooks, kits, and the like.

TABLE 2
PROPORTION OF EXPENDITURES FOR STANDARDIZED TESTS IN LARGE CITY, URBAN, SUBURBAN AND RURAL DISTRICTS 1975-76

District	Percent of Teaching Supplies	Percent of Total Operating Expense
A	2.6	.04
B	1.7	.02
C	1.5	.03
D	—	.08
E	—	.06
F	—	.05
G	—	.06
H	1.2	.02
I	1.5	.02
STATE est. $\bar{x}$	1.7	.03

The declines in pupil enrollments, and in the proportion of funds spent and available for new instructional materials are not the only declines we face. One of the most dramatic and serious is the decline in school achievement as evidenced by test scores and by evidence

of low functional literacy outside of the school environment. Test authors and publishers have know for some time of this general decline in student achievement. It became apparent for many of us as recently as seven years ago when new editions of standardized tests were being written and item analyzed prior to publication. In many instances and in some content areas, items which had functioned best at one grade level just eight years before now had psychometric qualities which dicated their movement upward by one or two grade levels. An item that may have functioned well at the fourth grade in the early 1960's now was too difficult for fourth graders and had to be moved up into the fifth or sixth grade if it was to be easy enough for the average child to respond correctly. Some of the reasons for this general decline have now surfaced. The most recent explanation for this has appeared in the October issue of *Phi Delta Kappan*.[4] In an interview with Willard Wirtz, chairman of an independent 21-member commission which investigated the decline in *SAT* scores, some reasons were identified as being increased television viewing, social unrest, curriculum instability, almost unrestricted diversification of learning activities based upon pupil choice, family disintegration, and the like.

I feel that interview was as noteworthy for what it did not say, as much as it was for what it did report. For example, even prior to the commission's attempts to examine the general test score decline, reports were available from international studies in Reading and Mathematics that a similar delcine in school achievement was noted in countries around the world. In fact, the slope of the decline reported in these studies was almost a carbon copy of the decline noted here in the United States. There are, of course, other equally plausible explanations for the decline in school achievement: the breakdown in respect for authority — statuatory or educational; massive infusions of curricular changes that were inadequately tested prior to extensive diffusion into the schools; the romanticizing of the inherent wisdom of youth and its corollary, the short-sightedness of the elderly (read that as anyone over 30). Unfortunately, however logical the explanations for the general decline in school achievement, the reasons for the decline are less important than the fact of the decline.

Paradoxically this drop in school achievement has brought about some of the increases of other forces which school administrators now confront.

The Increases

Foremost among these increases are (1) legislative demands for assessment and accountability, (2) bureaucracies and policy statements at state and national levels, (3) politicizing of testing by diverse groups, and a resultant intervention of the courts into the administration of the schools.

Legislative demands for accountability and assessment. At a recent conference called by the Education Commission of the States, material was presented which summarized state legislation concerning minimal competencies for high school graduation and grade-to-grade promotion.[5] The overall impact of reviewing the various pieces of legislation considered by the fifty states clearly indicates that we are on the upswing of a trend toward national, as well as state, concerns with the status of educational achievement in the United States.

As of June 1, 1977 twenty-four states had introduced 67 bills related to this concern. Twelve states had enacted legislation, 17 states had issued State Board of Education rulings supporting minimal comptency testing, six states had legislation pending, and eleven states had failed to enact legislation, evidence that should give scant comfort to those among us who are concerned about the intrusion of larger units of government into local affairs.

Not all of these legislative acts and state board of education policies dealt with the same topics. Some thought of minimal competencies as objectives to be reached at some time during the formal years of schooling; others distinguished between academic diplomas and certificates of attendance; others allowed students to leave school early if they could demonstrate adequate proficiency on a state mandated measure. Still others established key check points throughout the twelve-year period (a pupil could not progress, for example, from grades 3, 5, 8 and 11 unless he or she had passed a basic skills test). Some legislatures were concerned with survival or life-skills (the reading of labels on medicine bottles, the interpretation of a traffic ticket, the interpretation of a television program listing in a newspaper), while some legislation was geared specifically to compensatory aid based upon specific performance indicators. Buth whatever the reason served by the legislation, it is clear that there is an increasing concern in the land for the local schools tc demonstrate that the pupils enrolled are capable of profiting from the instruction given them.

Several bills have been introduced at the federal level, which if passed in part or in whole promise to give school superintendents a whole new area of anxieties to overcome. H.R. 6088 introduced in the House of Representatives April 5, 1977 calls for the establishment of basic educational standards in reading, writing, and mathematics, the establishment of a National Commission on Basic Education whose tasks have been defined as the establishment of those basic standards, the review of state plans to conform with the federal standards to established by the Commission, and the approval of state plans for funding under H.R. 6088. At the moment, this bill has been referred to the House Committee on Education and Labor. But observers of this type of legislation give it a good chance of passage in some form. Other federal bills have been introduced by Representatives Michael Harrington of Massachusetts and Albert Quie of Minnesota. These bills deal with the release of test data and with the finding of compensatory education programs.

Representative Quie, speaking before a testing conference, stated some beliefs which seem to encapsulate pretty well the rationale for much of the mandated assessments we now face. He said:

> We who are concerned with the schools and education must respond to:
>
> ... public dissatisfaction with achievment test scores,
> ... student unrest and rebellion,
> ... the public outrage ouver dropout rates and the number of functional illiterates graduating from our schools,
> ... the public intolerance of runaway costs (as manifested in increasing bond-issue failure, and premature school closings),
> ... and — especially in the minority-group sector — the increasing demands for community control, in the hope of that greater accountability can be brought about.
>
> The current push for greater accountability will *not* turn

out to be a passing fad. It is likely that we can expect increasing demands for accountability, and we can expect the educational establishment to respond positively to those demands and to resist them at the same time.

> . . . Assessment and the attendent accountability are a two-part phenomenon in which one party demands — or makes known his wish for — accountability, and the other party responds out of a sense of 'oughtness.' Accountability does not readily come if it is not asked for or mandated. So both parties will have to become educated — the one as to the ways and techniques of being accountable, and the other as to what can be expected in the way of accountability, the forms of mandating it, and the information that is required prior to asking for it. This is the human exchange in which we are now involved.[6]

Bureaucracies and policy statements. A fact of life is that state departments of education do not always wait for legislation to catch up with or authorize certain policies enunciated by them.

Of the fifty states reporting to the Education Commission of the States, it was noted that in seven states, assessment-related legislation had failed to pass the legislatures and yet the state boards of education or the departments of education in these states had operative policies which controlled some form of state assessment and accountability. As many as 15 states reported no legislation pending, passed, or failed which nevertheless had operative policies governing mandated assessment programs.

In the state of Michigan we are subjected to a large number of state board of education policies which govern the actions of the state department of education, and which have considerable impact upon local districts and classrooms. And while this is neither unheard of nor necessarily bad, it is often confusing. For in several instances the state educational bureaucracy has either outstripped authorizing legislation and has had to backtrack on its policy statements or else has issued conflicting policies which nullify other policies which had just been passed.

A good case in point occurred recently. Since 1969, the State Board of Education has developed an evolving program of assessment and accountability. It has moved from a traditional norm-referenced ranking of achievement by districts within the state to an objective-referenced, and now a norm-referenced ranking of the districts by objectives. Initially superintendents were told that there would be no use made of state assessment data for allocating state aid to districts. From the second year of the program however, compensatory funding has been based on pupil performance on state assessment measures.

In March of 1977 the State Board of Education approved a plan to field test a concept called, "GESTALT". GESTALT is an acronym for Getting Educational Specialists Thinking and Acting on Learning Theories. The plan has been a controversial one in Michigan since it was first introduced by the State Superintendent of Instruction in the 1976-77 school year. The plan is based on the theory that some school buildings need direct state help to meet individual needs. The Department of Education believes that it can best do this by working directly with a building principal and staff in both educational and fiscal matters. The State Board of Education at its March, 1977 meeting directed the Superintendent to seek amendments to Chapter 3 of the State School Aid Act for 1977-78. The amendments were designed to facilitate the implementation of GESTALT.

In order to implement GESTALT, the State Board of Education has set criteria based on the Michigan Educational Assessment Program (MEAP) to identify "high needs" schools. Specifically the criteria are these:

1. Minimum expectation for individual student achievement will be attainment of 75% or more of the math and reading objectives measured by MEAP.
2. If 50% or more of the students have not attained 75% of reading and/or mathematics objectives, the school will be identified as a "high needs" school.
3. Based on identification data from the last three years, a school would be considered "high needs" if it met the criteria for two of the last three years.
4. Trends will be identified as "improving" when a school shows a year-to-year increase of 5% in the number of students attaining 75-100% of the reading and math objectives and a 5% decrease in the number attaining 0-24% of the objectives.[7]

But, just six months later on September 7, 1977 the State Board of Education passed a resolution on the state assessment program which reads in part,

> . . . be it
> RESOLVED, That the State Board of Education reaffirm its position that state assessment data do not portray the total school district or school building performance and, therefore, should never be used as the evaluation of teachers, school buildings, or school districts, individually or in a comparative nature; and be it finally
> RESOLVED, That this resolution be placed on the bulletin board of all elementary and secondary schools in the state and be disseminated to all major news media sources.[8]

So, where does that leave the administrators of local districts and, for that matter, the staff of the state department of education charged with implementing project GESTALT? I don't know, and neither do some high administrators in the Michigan Department of Education.

Of course this latest incident is not the first one which has caught either the Department of Education or the State Board of Education off-base. Just one month before school was to start this current academic year, the legislature refused to fund a 10th grade assessment which the State Superintendent had assured the local school districts would take place. In fact, all the 10th grade instruments had been construced and plates prepared when the word came back saying that enabling legislation had been refused because the legislature was not convinced that the data so obtained would be used meaningfully enough to warrant the expenditures necessary to implement the program.

But lest the reader think that this paper is related to Michigan's assessment program only I would turn for a moment to New York. It would seem to me, from information given me by some New York administrators, that similar problems of increasing bureaucratic policies are emerging in that state as well. It appears that their Pupil Evaluation Program (PEP), Programs for Pupils with Special Educational Needs (PSEN), and the various related state-aid policies potentially carry the seeds of similar misunderstandings. Some administrators have indicated a concern for an overreliance on the state's

Basic Competency Tests being tied to graduation requirements. There has been a concern voiced by some New York administrators that PSEN is tied to state aid through Chapter 241. These concerns may not necessarily be serious criticisms . . . , but then again, they might be. What is certain is that anytime a remote office intrudes upon the daily activities of a local enterprise, misunderstandings will develop for that is one of the serious faults of any state mandated program.

An examination of some correspondence and other local school records indicates that New York teachers and administrators are really no different than those of other states. They evidence a concern that state mandated tests will be used to judge them unfairly because of the performance of pupils in their rooms. They feel, with some justification, that output information in the form of achievement tests do not explain why a pupil scores as he or she does, but only indicate the status of that learning. And this brings us to a third and final comment about the increase in certain demands made upon the administrator, *i.e.*, the politicizing of test results.

The politicizing of testing. It has become chic to disparage scholastic aptitude measures as bearers of evil tidings. The courts have now before them law suits which allege that scholastic aptitude measures are used to systematically exclude minority children from regular classroom instruction. Such a case is pending before a San Francisco court. Lawyers for the NAACP predict that if the case is decided in their favor, the ripple effect of that decision will cause wholesale voluntary abandonment of the use of these tests. I believe that their assessment is probably correct. And I also believe that it would be unfortunate if such a decision is rendered and if such an effect takes place.

There is a far greater amount of emotional involvement on the parts of administrators, board members, teachers, and psychologists concerning the use of scholastic aptitude tests than is warranted given the intellectual training they posses. Any issue as serious as those raised by opponents of this type of test should be subjected to a rational analysis rather than the emotion we see. Furthermore, what is often at stake here is not an educational issue but a political one that appeals, at times, to the worst in us.

The Michigan Department of Education at one time used a scholastic aptitude measure taken directly from a well-known and widely marketed test. They referred to it, in the initial years of their state assessment program, as a vocabulary test, and treated it as an output measure. They have since discontinued that use after they received considerable criticism for using the test in that way. The concern of the critics did not lie in the use of a scholastic aptitude measure, but that it was misused and mislabeled for political reasons. The Department knew that if they were to identify it for what it was, they would be subjected to considerable pressure from political action groups opposed to this type of test. The result? No real measure of input that could be used to help administrators interpret the output measures to their constituencies. New York used an approach called PIE (Performance Indicators in Education), to provide some input information to help administrators in their interpretation of the PEP tests. But it is interesting to note that even though they identify three factors, families whose primary means of support is a public welfare program, enrollment, and heterogeneity of pupil PEP scores, only one measure is given credence — public welfare program support.[9] The expectancy measure resulting from the application of this variable is certainly more helpful than the vacuum that Michigan administrators work in, but isn't this also an indication of a sensitivity to the political impact of using a scholastic aptitude measure that has proven itself over the years to be at least as sensitive to a child's propensity to handle school-related learning as any socio-economic measure?

The pity of it all is that we stand the danger of giving up a useful measure because it has been made a political *cause celebre*.

Fitzgibbon, in an address to the American Personnel and Guidance Association, in 1973, discussed the politics of testing. He made one point that is especially relevant to these remarks. He observed that:

> *Political pressures are causing some people to spend more time changing or charging the test than in trying to help the pupil who scores low.* If the test is inadequate, it should of course be removed from use. My observation, however, is not about that relatively clear-cut situation. Rather, it is about one in which the test is satisfactory but yields results "unacceptable" on the political front. Here we see charges that the norms are unfair or poorly constructed; that the questions are biased or, at best, irrevelant; or simply that the complainants didn't want to give the test in the first place. Here too we find the test attacked, or attempts made to modify it, as fear of the impact of poor results upon the community grows. This fear can become debilitating as well as dangerous to sound testing procedures.[10]

Interpreting and Using The Data We Collect[11]

Up until this point this paper has addressed itself to those factors which turn gentle superintendents into hyperactive, unusually sensitive human beings. We have just barely touched upon the declines that trouble us and the increase in some pressures that plauge us. And now I turn to the last of the three points that I believe are relevant to this topic, *viz.*, the means we employ to use and interpret the test information we collect.

A major characteristic and at the same time a possible fault of many accountability programs is the primary emphasis upon pupil's scores on standardized achievement tests. From these scores, newspapers and other mass media have either praised or condemned the effectiveness of teachers and of the educational programs which they have produced. But the mass media cannot always be faulted; too often school district officials and some state departments of education have released test data without adequate elaboration or without emphasizing the relationship between achievement test scores and other, equally important data. The result has often been that School A is reported as having achieved a group stanine of 7 in Total Reading, while School B is reported as having achieved a group stanine of 5. This then places School B in the awkward position of having to explain why it is "not as effective" as School A. What may be missing from these reports, and it all too often is, is that School B has a stanine of 4 in the scholastic aptitude measure. Now it would appear that School B may be more effective. But there are still other bits of data which need to be collected before one can come to any meaningful conclusion.

If a district has not developed an educational auditing system which collects some key baseline data, it would be advisable, in the light of contemporary accountability systems, for it to do so. This might include the following

types of information:

1. Occupational level of parents, using current census data categories.
2. Median years of schooling of adults 25 years old and over.
3. Percent of pupils who received their education within the district for all years preceding each grade level tested (an index of mobility).
4. Percent of teachers holding graduate degrees.
5. Median years of professional experience of teachers.
6. Professional staff/pupil ratio.
7. Teacher/pupil ratio.
8. Other data that have local significance.

These data could be collected by classroom, building, or system and could be reported whenever public release of test data is contemplated. These types of information will enable the school district, the mass media, and the public to evaluate more meaningfully the results of standardized achievement and scholastic aptitude testing.

All too often, test data are analyzed and used to make decisions about the effectiveness of school programs and staff in light of the test results alone. This quantum leap in judgment from only one set of input variables (teacher efforts) to an output measure (test scores) suffers from oversimplification and faulty logic. Such a procedure assumes that the characteristics of all children are the same, and that the homes and communities in which they live do not differ from one another. In truth, as we know, families do differ considerably; children do not possess the same propensities for school-related learning; communities do differ in their ability to offer auxilliary services that could enhance the school learning of pupils; stability of school attendance is varied; and the attitude of pupils and their peers towards school achievement does differ considerably from community to community and from neighborhood to neighborhood. Thus to heap praise upon some school programs and school teachers solely for high achievement of pupils is as faulty as it is to condemn other schools and teachers for low performance on standardized achievement tests.

It seems to me that three important variables must be considered when making judgements concerning the effectiveness of teachers and/or school programs:

1. the characteristics of the pupils being taught,
2. the characteristics of the homes and communities from which these pupils come, and
3. those school related factors which may have an effect on test scores.

The total influence of this complex of variables plus the child's health and nutrition are what, in large measure, determine pupil performance on standardized tests.

Much has been said of the first two characteristics and I doubt that I could add anything new to the dialogue. But what of the third characteristic, *viz.*, school-related factors? While we cannot ignore the non-school factors which contribute to the level of achievement of a pupil, neither can we ignore those variables within the school which bear upon academic performance. While recent research would indicate less of a relationship between achievement and certain cost-related factors than was once supposed, factors including class size, the availability of auxilliary services such as counseling and guidance programs, remedial programs, extensive libraries, and the experience and annual turn-over rate of teachers show a low but positive relationship to achievement.

REPORTS AVAILABLE

Back issues at *Measurement in Education* are available for 1.50 each per single issue, and for 75¢ each in quantities of 25 or more for a single issue.

Vol. 1,	No. 1	*Helping Teachers Use Tests* by Robert L. Thorndike
	No. 2	*Interpreting Achievement Profiles — Uses and Warnings* by Eric F. Gardner
	No. 3	*Mastery Learning and Mastery Testing* by Samuel T. Mayo
	No. 4	*On Reporting Test Results to Community Groups* by Alden W. Badal & Edwin P. Larsen
Vol. 2,	No. 1	*National Assessment Says* by Frank B. Womer
	No. 2	*The PLAN System for Individualizing Education* by John C. Flanagan
	No. 3	*Measurement Aspects of Performance Contracting* by Richard E. Schutz
	No. 4	*The History of Grading Practices* by Louise Witmer Cureton
Vol. 3,	No. 1	*Using Your Achievement Test Score Reports* by Edwin Gary Joselyn & Jack C. Merwin
	No. 2	*An Item Analysis Service for Teachers* by Willard G. Warrington
	No. 3	*On the Reliability of Ratings of Essay Examinations* by William E. Coffman
	No. 4	*Criterion-Referenced Testing in the Classroom* by Peter W. Airasian and George F. Madaus
Vol. 4,	No. 1	*Goals and Objectives in Planning and Evaluation: A Second Generation* by Victor W. Doherty and Walter E. Hathaway
	No. 2	*Career Maturity* by John O. Crites
	No. 3	*Assessing Educational Achievement in the Affective Domain* by Ralph W. Tyler
	No. 4	*The National Test-Equating Study in Reading (The Anchor Test Study)* by Richard M. Jaeger
Vol. 5,	No. 1	*The Tangled Web* by Fred F. Harcleroad
	No. 2	*A Moratorium? What Kind?* by William E. Coffman
	No. 3	*Evaluators, Educators, and the Publics: A Detente?* by William A. Mehrens
	No. 4	*Shall We Get Rid of Grades?* by Robert L. Ebel
Vol. 6,	No. 1	*On Evaluating a Project: Some Practical Suggestions* by John W. Wick
	No. 2	*Measuring Reading Achievement: A Case for Criterion — Referenced Testing and Accountability* by S. Jay Samuels and Glenace E. Edwall
	No. 3	*Security in a Citywide Testing Program* by Anthony J. Polemeni
	No. 4	*Dear Mama: Why don't they love me anymore?* by Thomas J. Fitzgibbon
Vol. 7,	No. 1	*The Discrepancy Evaluation Model (Part I)* by Andres Steinmetz
	No. 2	*The Discrepancy Evaluation Model (Part II)* by Andres Steinmetz
	No. 3	*What Graduate and Professional School Students Think About Admission Tests* by Letnard L. Baird
	No. 4	*The Paradox of Education Testing* by Robert L. Ebel
Vol. 8,	No. 1	*FACE UP: A Framework for Assessing Career Educators Utilization Programs* by Bruce W. Tuckman
	No. 2	*Grade Equivalent Scores* by Gary Echternacht and *If Not Grade Equivalent Scores — Then What?* by Jeanne M. Plas
	No. 3	*Criticisms of Testing: How Mean is the Median?* by Henry S. Dyer

Outmoded instructional materials, lack of adequate laboratory equipment, and poor facilities for learning are all factors that can affect school achievement.

Some schools have a relatively high proportion of non-English speaking pupils. If the school has not prepared an instructional program that takes this kind of pupil population into account, its academic performance on achievement tests may yield spurious results.

As children progress through school, they encounter and must be affected by many different people who work in the school. Some will be sensitive to each child's needs, others will not. Over the years the child is likely to be influenced by administrators, special teachers, custodians, secretaries, para-professionals, and other pupils as well as by classroom teachers. The sum total of experiences with all of these people plus interaction with the family, the neighborhood, and the child's inherent capabilities determine in large measure performance on a standardized achievement test. The score that one observes on such a test, then, is the result of the interaction of all of these variables. One cannot assign fault or praise to any one of these variables and identify it as the sole contributor to the score.

Superintendents would be well-advised to caution the community that education is cumulative and the result of all of the forces acting upon a pupil. The essential thing is not, as some accountability programs stress, to assign blame or credit to any one person or program; it is understanding what a child does indeed know and then proceeding to build a program that will enable him to become a literate adult. Schools must be prepared to accept the fact that there are factors which contribute to pupil achievement of school learning over which the school has little or no control.

In Summary

I have tried in this presentation to trace the milieu of the superintendent and relate that milieu of declining enrollments, resources, test scors and the like, to the increased pressures he or she faces. And while they are connected, I would hope that we do not lose our perspective in the light of what seems to be decreasing control of our local school districts.

I hope that each of us can provide our school districts with some added insights, some new concepts to consider, and most certainly a reinforcement that what we are doing for the children and adults in our community is important. And I hope above all else that we become more convinced than ever that while we must be sensitive to the data we receive locally and from the state concerning pupil progress that we not consider ourselves, our staffs and our programs as failures if some portion of our children do not achieve as well as we would like them to do. For we are educationists and not miracle workers. The home, the community and the school share in the intellectual and social development of our young. There is no meaningful way to disaggregate that responsibility and to assign credit or blame to any one component.

1. Stanley E. Hecker and Fred R. Ignatovich. *Michigan Total School Members Projection Based On 1971-72 Through 1976-77 Membership Data* (East Lansing, Michigan: Department of Administration and Higher Education, Michigan State University, September 27, 1977, mimeographed, 10 pp.).
   Figures 1 and 2 were taken from this report to the Michigan Council of State College Presidents.
2. Data for Table 1 were drawn from annual state financial summaries reported by the Michigan Department of Education for the years 1971-72 to 1975-76.
3. This sample, while not randomly drawn, is a good representation of size of population centers represented in the state summaries in Table 1.
4. Harold G. Shane. "The Academic Score Decline: Are Facts the Enemy of Truth? — An Interview with W. Willard Wirtz." *Phi Delta Kappan* 59:83-86, 145-146 (October 1977).
5. Chris Pipho. *Update VI: Minimal Competency Testing.* (Denver, Colorado: Education Commission of the States, June 1, 1977, mimeographed, 48 pp.).
6. Albert H. Quie. "An Elected Official Looks at Mandated Assessment" in *Focus On Evaluation: Mandated Assessment* (New York: Harcourt Brace Jovanovich, 1974) p. 4.
7. *Needs Criteria For Michigan Schools* (Lansing, Michigan: Michigan Department of Education, 1977) p. 3.
8. Taken from a Michigan State Board of Education Resolution, "State Assessment Program" adapted September 7, 1977.
9. *Performance Indicators in Education* (Albany, New York: The State Education Department, Bureau of School and Categorical Programs Evaluation, Spring 1977, mimeographed, 10 pp.).
10. Thomas J. Fitzgibbon. *Focus on Evaluation: Political Use of Educational Test Results.* (New York: Harcourt Brace Jovanovich, 1973) p. 4.
11. The material which follows was adapted from Richard Madden, Eric F. Gardner, Herbert C. Rudman, Bjorn Karlsen and Jack C. Merwin. *Administrator's Guide for Interpreting Standford Achievement Test* (New York: Harcourt Brace Jovanovich, 1974).

NCME NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION

202 S. Kedzie Hall
Michigan State University
East Lansing, Michigan 48824

Second class postage paid at East Lansing, Mich. 48823 and additional mailing points.